Project Name
Apartment interest level prediction analysis

Target Release	Nov 28th, 2022
Document Status	Draft
Team Number	Group 4
Team Lead	
Team Members	Miao Yuan, Wenrui Fang, Yan Yan, Shan Xu

🎯Objective

Provide context on this project and explain how it fits into your organization's strategic goals.

The factors that influence the rental market, the home buying and selling market are similar, but there are some differences. Relatively speaking, rental housing is less affected by the macro-economy, more affected by supply and demand, and relatively flat in price fluctuations. We may have the following questions about house rental: What is the distribution of rental housing in different cities? What is the impact of various business districts in the city on rents? How does traffic affect rent? What is the impact of room type (apartment or house) on rent? How does the degree of newness (hardcover or not) affect the price? Which apartment types are popular? What do renters pay most attention to? These things can be dug up. So, for this project, we want to build a model that predicts people's preference for different houses when renting. We will use Python 3 in Jupyter Notebook environment and plan to use pandas, numpy, matplotlib, scipy, sklearn, seaborn, xgboost, etc.

📊 Success metrics

List project goals and the metrics you'll use to judge its success.

Goal	Metric

Hit each milestone on time	0% late submission
Maximize accuracy by applying different models	maximize the accuracy
Employee satisfaction	100% Employee Net Promoter Score

🤔Assumptions

We assume the factors that influence the rental market and the home buying and selling market are similar, but there are some differences.

Get data from Kaggle competition which has data labeled in a consistent way.

🌟Milestones

Create a visual roadmap and help your team stay on track.

- Estimated time for planning: 1 week
- Estimated time for pre-processing data: 0.5 week
- Estimated time for applying models and analyzing data: 2.5 weeks
- Estimated time for testing and evaluation: 1 week
- Estimated time for drawing conclusions: 1 week

🗒️Requirements

Requirement	User Story (Details)	Importance

Evaluate data entry attributes, weed out least relevantly least necessary attributes	Factors with more weight for example: - number of bathrooms, number of bedrooms, location, price, Factors with less weight for example: - built date, leasing managers, photo quality	**High**
Apply multiple models to do prediction	For example: XGBoosting Model, random forest and Ada boost, etc.	**High**
Model prediction performance comparison	Compare the prediction performance conducted from different models	**High**

❓Open Questions

- Evaluate different data sources and stick with the most appropriate one
- Think about how to select proper attributes?
- How to determine if we have enough data?
- Which models are appropriate to use on our data?
- How can the model be maintained over time?

⚠️Out of Scope

- Redesign the recommendation system and receive responses of interest from customers.
- Use deep learning models like Convolutional Neural Networks (CNNs) to train our data and analyze.